

2. uzdevums

Doti teikumi latviešu valodā un to tulkojumi uz dirbalu valodu jauktā secībā:

- | | |
|--------------------------------------|--|
| 1. Jūs atnācāt un piekāvēt mūs. | A. nana baniju purra balgan |
| 2. Jūs piekāvēt mūs, un mēs atnācām. | B. nana purrana balgan nana baniju |
| 3. Mēs piekāvēm jūs un atnācām. | C. purra baniju purra nanana balgan |
| 4. Mēs atnācām un jūs piekāvēt mūs. | D. purra nanana balgan baniju |

1. uzdevums Savietojiet teikumus ar pareizajiem tulkojumiem!

Nav grūti pamanīt, ka **baniju** un **balgan** ir darbības vārdi, bet **nana**, **nanana**, **purra** un **purrana** – vietniekvārdi. Taču, saskaitot vietniekvārdus latviešu un dirbalu teikumos, saskaramies ar problēmām.

Mēs – 3, mūs – 3 (kopā 6). nana – 3, nanana – 2 (kopā 5).
Jūs – 3, jūs (akuzatīvā) – 1 (kopā 4). purra – 4, purrana – 1 (kopā 5).

Tas var rosināt uz domu, ka dirbalu valodā vietniekvārdus izmanto citādāk, nekā latviešu valodā. Latviešu valodā ir iespējams izlaist vienādus vietniekvārdus, pārvēršot saliktu sakārtotu teikumu par vienkāršu teikumu ar vienlīdzīgiem teikuma locekļiem
Piemēram, “jūs atnācāt, un jūs piekāvēt mūs” vietā sakām “jūs atnācāt un piekāvēt mūs”.

Rekonstruēsim izlaistos vietniekvārdus latviešu teikumos:

Jūs atnācāt un JŪS piekāvēt mūs.
Jūs piekāvēt mŪS, un mēs atnācām.
Mēs piekāvēm jūs, un MĒS atnācām.
Mēs atnācām, un jūs piekāvēt mūs.

No jauna skaitam vietniekvārdus:

Mēs – 4, mūs – 3 (kopā 7). nana – 3, nanana – 2 (kopā 5).
Jūs – 4, jūs (akuz.) – 1 (kopā 5). purra – 4, purrana – 1 (kopā 5).

Tagad sanāk, ka latviešu teikumos ir 12 vietniekvārdi, bet dirbalu teikumos – 10. Tātad, varam pieļaut, ka dirbalu valodā arī ir iespējams izlaist kādus vietniekvārdus, tikai ne tos, kurus latviešu valodā.

Apskatoties uz dažādu vietniekvārdu skaitu abās valodās, varam pamanīt, ka “jūs” un “jūs (akuz.)” ir sastopami tik pat daudz reižu, cik **purra** un **purrana**. Tad var pieļaut, ka “jūs” ir **purra**, bet “jūs (akuz.)” – **purrana**. Tādā gadījumā, “mēs” būtu **nana**, bet “mūs” – **nanana**. Sanāk, ka pa reizei tiek izlaisti vietniekvārdi “mēs” un “mūs”, taču ar to tiksīm skaidrībā vēlāk, jo mums ir gana informācijas, lai pagaidām atbildētu vismaz uz 1. jautājumu.

Jautājums 1. Savietojiet teikumus ar pareizajiem tulkojumiem!

Jūs atnācāt un piekāvāt mūs.	=	purra banīpu purra nanana balgan (burtiski: jūs atnācāt jūs mūs piekāvāt)
Jūs piekāvāt mūs, un mēs atnācām.	=	purra nanana balgan banīpu (burtiski jūs mūs piekāvāt atnācām)
Mēs piekāvām jūs un atnācām.	=	nana purrana balgan nana banīpu (burtiski mēs jūs piekāvām mēs atnācām)
Mēs atnācām, un jūs piekāvāt mūs.	=	nana banīpu purra balgan (burtiski mēs atnācām jūs piekāvāt)

2. jautājums Pārtulkojiet uz latviešu valodu:

nana purrana balgan banīpu burtiski var iztulkot kā **mēs jūs piekāvām atnācām(t)**.

Pagaidām vēl nav skaidrs, kurš atnāca: mēs atnācām vai jūs atnācāt. Par to, ka pareizā atbilde ir tieši “mēs jūs piekāvām, un jūs atnācāt”, netieši liecina tas, ka teikums “mēs jūs piekāvām un atnācām” jau ir dots un dirbalu valodā izskatās citādi.

Taču šis pieņēmums ir vēl jāpārbauda. Paskatīsimies uz to, kuri vietniekvārdi tiek izlaisti. Teikuma “jūs piekāvāt mūs, un mēs atnācām” tulkojumā ir izlaists “mēs”, bet teikuma “mēs atnācām, un jūs piekāvāt mūs” tulkojumā izlaists “mūs”.

Mēs (**nana**) ir izlaists teikumā, kur jau bija pieminēts vārds “mūs” (**nanana**), bet “mūs” izlaists teikumā, kur jau bija pieminēts vārds “mēs”.

Tad sanāk, ka **ja vienas teikuma daļas objekts ir izteikts ar to pašu vietniekvārdu kā citas subjekts (jeb priekšmets), tad otrais no vietniekvārdiem tiek izlaists**.

Tādā gadījumā arī pieņēmums, ka teikumā **nana purrana balgan banīpu** (burtiski – “mēs jūs piekāvām atnācāt”) izlaists tieši vārds “jūs”, ir apstiprināts, jo pirmajā teikuma daļā jau ir “jūs (akuz.)”.

nana purrana balgan banīpu = Mēs piekāvām jūs (akuz.), un jūs atnācāt.

Uzdevumā ir ilustrēta t.s. sintaktiskās ergativitātes parādība. Dirbalu valodā ir iespējams izlaist vietniekvārdus šādā veidā, jo, atšķirībā no latviešu valodas, kur teikuma ar netrāsītīvu darbības vārdu (piem., “mēs atnācām”) subjekts tiek grupēts (un citādi uzvedas līdzīgā veidā) ar teikuma ar trāsītīvu darbības vārdu (piem., “mēs piekāvām jūs”) subjektu, bet valodās, kurās izpaužas sintaktiskā ergativitāte, teikuma ar netrāsītīvu darbības vārdu subjekts tiek grupēts ar teikuma ar trāsītīvu darbības vārdu objektu.

Citiem vārdiem, teikums, kurš izskatās kā “Es biju ieraudzījis Andri un pateicis “Sveiks!” ” latviski nozīmē “Es biju ieraudzījis Andri un es biju pateicis “Sveiks!” ”, bet dirbalu valodā nozīmētu “Es biju ieraudzījis Andri un viņš bija pateicis “Sveiks!” ”.

Tagad mums pietiek informācijas, lai atbildētu arī uz 3. jautājumu.

Jūs atnācāt, un mēs piekāvām jūs (akuz.). = **purra banīpu nana balgan**
(burtiski **jūs atnācāt mēs piekāvām**)

Jūs piekāvāt mūs un atnācāt. = **purra nanana balgan purra banīpu**
(burtiski **jūs mūs piekāvāt jūs atnācāt**)

3. uzdevums

Doti datumi solresol valodā un to tulkojumi jauktā secībā:

- | | |
|-----------------------------------|-------------------------------|
| 1. solsolfa, reremi mimifa | A. sestdiena, 20. janvāris |
| 2. solsolfa, reresi mifafa | B. trešdiena, 3. maijs |
| 3. solsolmi, dolala refafa | C. trešdiena, 8. augusts |
| 4. solsolmi, reremi mimire | D. ceturtdiena, 9. augusts |
| 5. solsolsi, dorere fafare | E. sestdiena, 17. augusts |
| 6. solsolsi, reremi milala | F. ceturtdiena, 15. decembris |

1. jautājums Savietojiet datumus ar pareizajiem tulkojumiem!

Trīs reizes ir sastopams vārds ‘**reremi**’, un trīs reizes ir sastopams vārds “augusts”, no kā arī veicam attiecīgu secinājumu.

Mums ir 8., 9. un 17. augusts, kuriem atbilst **reremi mimire**, **reremi mimifa** un **reremi milala**. Ja pieņem, ka vārdi, kuri atšķiras tikai ar vienu noti (**mimire** un **mimifa**) – ir 8 un 9, tad pareizās atbilstības var noskaidrot ātri.

- | | |
|-----------------------------------|-------------------------------|
| 1. solsolfa, reremi mimifa | D. ceturtdiena, 9. augusts |
| 2. solsolfa, reresi mifafa | F. ceturtdiena, 15. decembris |
| 3. solsolmi, dolala refafa | B. trešdiena, 3. maijs |
| 4. solsolmi, reremi mimire | C. trešdiena, 8. augusts |
| 5. solsolsi, dorere fafare | A. sestdiena, 20. janvāris |
| 6. solsolsi, reremi milala | E. sestdiena, 17. augusts |

Lai atrisinātu pārējos uzdevumus, sastādīsim tabulu no esošajiem mēnešiem un nedēļas dienām.

Jan	dorere		Mon	
Feb			Tue	
Mar			Wed	solsolmi
Apr			Thu	solsolfa
May	dolala		Fri	
Jun			Sat	solsolsi
Jul			Sun	
Aug	reremi			
Sep				
Oct				
Nov				
Dec	resesi			

Var redzēt, ka mēnešu un dienu nosaukumi sastāv no divām daļām: “**nots1-nots2nots2**” vai “**nots1nots1-nots2**”.

Turklāt, jo augstākas ir **nots1** un **nots2** gammā **do-re-mi-fa-sol-la-si**, jo vēlāks ir mēnesis.

Turklāt noti2 paaugstina ar soli 1.

Aizpildīsim tukšās vietas:

Jan	dorere		Mon	solsoldo
Feb	domimi		Tue	solsolre
Mar	dofafa		Wed	solsolmi
Apr	dosolsol		Thu	solsolfa
May	dolala		Fri	
Jun	dosisi		Sat	solsolsi
Jul			Sun	
Aug	reremi			
Sep	rerefa			
Oct	eresol			
Nov	rerefa			
Dec	reresi			

Jūlija vietā der gan **reredo**, gan **rerere**, bet piektdienas vietā – gan **solsolsol**, gan **solsolla**. Loģiski ir pieņemt, ka trīs identisku nošu kombinācija (citiem vārdiem, gadījums, kad nots1 = nots2) tiek izlaista. Šo hipotēzi apliecina arī novērojums, ka mēneši sākas ar **dorere**, nevis ***dododo**.

Jan	dorere		Mon	solsoldo
Feb	domimi		Tue	solsolre
Mar	dofafa		Wed	solsolmi
Apr	dosolsol		Thu	solsolfa
May	dolala		Fri	solsolla
Jun	dosisi		Sat	solsolsi
Jul	reredo		Sun	?
Aug	reremi			
Sep	rerefa			
Oct	eresol			
Nov	rerefa			
Dec	reresi			

Pagaidām nav līdz galam skaidrs, ko darīt ar svētdienu; to atstāsim vēlākam.

Izrakstīsim visus dotos skaitļa vārdus.

1			11	
2			12	
3	re-fafa		13	
4			14	
5			15	mi-fafa
6			16	
7			17	mi-lala
8	mimi-re		18	
9	mimi-fa		19	
10			20	fafa-re

Loģiski pieņemt, ka skaitļu vārdi veidojas līdzīgā veidā, kā dienas un mēneši.

1	<i>re-dodo</i>		11	<i>mimi-la</i>
2	<i>re-mimi</i>		12	<i>mimi-si</i>
3	re-fafa		13	<i>mi-dodo</i>
4	<i>re-solsol</i>		14	<i>mi-rere</i>
5	<i>re-lala</i>		15	mi-fafa
6	<i>re-sisi</i>		16	<i>mi-solsol</i>
7	<i>mimi-do</i>		17	mi-lala
8	mimi-re		18	<i>mi-sisi</i>
9	mimi-fa		19	<i>fafa-do</i>
10	<i>mimi-sol</i>		20	fafa-re

No šejienes skaidrāk redzams, ka skaitļa vārdus veido no cikliem garumā 12, ar apakšcikliem garumā 6: nots1 vispirms sešas reizes parādās dubultotā veidā, un pēc tam – sešas reizes nedubultotā veidā, pēc kā to pašu atkārti ar nākamo noti l.

Tagad varam atgriezties pie svētdienas.

Mon	<i>solsoldo</i>
Tue	<i>solsolre</i>
Wed	solsolmi
Thu	solsolfa
Fri	<i>solsolla</i>
Sat	solsolsi
Sun	<i>soldodo</i>

Tā kā nots1 vispirms parādās dubultotā veidā, un pēc tam nedubultotā, un ne otrādi, svētdiena ir **soldodo** (nevis ***ladodo**, kā varētu kļūdaini iedomāties).

Tagad mūsu rīcībā ir visa informācija, lai izpildītu pārējos uzdevumus.

2. jautājums

7. **solsolre, reresol resolsol**
8. **solsolla, domimi midodo**

otrdiena, 4. oktobris
piektdiena, 13. februāris

3. jautājums

9. pirmdiena, 6. jūlijs
10. svētdiena, 19. marts

solsoldo, reredo resisi
soldodo, dofafa fafado

4. uzdevums

Automātiskajā dabiskās valodas apstrādē (piemēram, veicot automātisku ortogrāfijas pārbaudi) bieži ir jānosaka, cik dažādi ir divi uzrakstīti vārdi. Viens no šim nolūkam izmantojamajiem kvantitatīvajiem mēriem ir tā saucamais “Damerau – Levenšteina attālums”

Šis attālums nosaukts par godu Vladimiram Levenšteinam un Frederikam Damerau. Levenšteins izdomāja veidu, kā mērīt “attālumus” starp vārdiem, bet Damerau neatkarīgi no Levenšteina piedāvāja veidu, kā klasificēt drukas kļūdas vairākās kategorijās.

Doti angļu vārdu pāri un Damerau – Levenšteina attālumi starp katra pāra vārdiem. Daži skaitļi ir izlaisti.

Vārdu pāris		Damerau – Levenšteina attālums
1. acre	car	2
2. anteater	theatre	4
3. banana	nanny	3
4. cat	crate	2
5. cocoon	cuckoo	3
6. emporium	empower	4
7. goer	ogre	2
8. lyra	lay	2
9. life	death	5
10. point	sirloin	5
11. stone	sonnet	3
12. surge	ruse	3
13. task	tusk	1
14. ape	ea	
15. baba	arab	
16. contest	toner	
17. eel	lee	
18. martial	marital	
19. monarchy	democracy	
20. seatback	backseat	
21. warfare	farewell	
22. peat	tape	
23. smoking	hospital	

1. jautājums. Aizpildiet tukšās vietas!

2. jautājums. Piedāvāriet definīciju Damerau – Levenšteina attālumam, un izdariet pieņēmumu, kādās kategorijās Damerau piedāvāja iedalīt drukas kļūdas!

3. jautājums. Doti divi vārdi ar garumiem m un n ($m > n$). Kāds ir maksimālais iespējamais Damerau – Levenšteina attālums starp šiem vārdiem? Kāds ir minimālais attālums? (Izsakiet atbildes izmantojot m un n).

Sāksim ar 2. uzdevumu. Damerau – Levenšteina attālums starp diviem vārdiem ir minimālais elementāro operāciju skaits, kas ir jāveic, lai iegūtu vienu vārdu no otra. Elementārās operācijas ir sekojošas: viena simbola izdzēšana; viena simbola pievienošana; viena simbola aizstāšana ar otru; divu blakusesošu simbolu maiņa vietām.

Attiecīgi arī četras kategorijas, kurās Damerau iedalīja drukas kļūdas, izstrādājot to automātiskās labošanas algoritmu, ir: simbola izlaišana; lieks simbols; simbola aizstāšana ar otru; divu blakusesošu simbolu sajaukta secība.

Dažreiz dalībnieki atbildot uz 2. jautājumu centās izklāstīt detalizētu algoritmu, kas ļautu atrast Damerau – Levenšteina attālumu. Taču tas ir diezgan komplicēti, un arī netika prasīts uzdevumā. Pietika ar definīciju, kura būtu līdzīga definīcijai iepriekšējā rindkopā.

Tagad varam atbildēt uz 1. jautājumu:

	Vārdu pāris		Damerau – Levenšteina attālums
15.	baba	arab	3
16.	contest	toner	4
17.	eel	lee	2
18.	martial	marital	1
19.	monarchy	democracy	5
20.	seatback	backseat	8
21.	warfare	farewell	6
22.	smoking	hospital	7
23.	ape	ea	2

Pie dažiem pāriem varam pakavēties atsevišķi. Tā kā vietām var mainīt tikai **blakusesošus** simbolus (tas izriet no dotā *peat–tape* pāra, kurai attālums ir 4, nevis 2), attālums starp *eel* un *lee* ir 2, nevis 1.

Tā kā visas operācijas notiek ar atsevišķiem simboliem, nevis simbolu grupām (kas arī ir redzams no pāra *peat–tape*), attālums starp *seatback* un *backseat* ir nevis 1 (kā intuitīvi varētu šķist), un nevis 4, kā nereti atbildēja dalībnieki, bet 8.

3. jautājums

Saprotams, ka Damerau – Levenšteina attālums nevar pārsniegt garākā vārdā garumu (aizstājot tajā simbolus, no tā var iegūt pilnīgi jebko), tāpēc maksimālais attālums ir **m**. Minimālais attālums ir **m – n** (simbolu skaits, kurš jāpievieno īsākajam no vārdiem, lai dabūtu garāko). Daudzi atbildēja 1, taču šāda atbilde neder, jo tika prasīts izteikt atbildi caur **m** un **n**.